基于专利的企业潜在研发伙伴推荐方法研究

翟东升　郭　程　张　杰　夏　军
(北京工业大学经济与管理学院　北京 100124)

摘要: 【目的】针对如何准确识别潜在研发伙伴, 提出一种基于专利的企业潜在研发伙伴的推荐方法。【方法】以 TRIZ 理论为基础, 对相关专利的功能、科学效应、功效等语义特征进行抽取, 构建出领域专利技术树, 将企业技术需求中的关键信息进行抽取并匹配到技术树中, 根据专利权人得出潜在研发伙伴并使用层次分析法对其进行评估。【结果】研究获取热水器防水垢技术领域的德温特专利数据, 对潜在研发伙伴进行识别与评估, 证明该方法的可行性。【局限】对于语义特征的抽取, 由于中文语法结构存在较大的灵活性, 抽取的准确率有待提高。【结论】该方法可以发现并评估潜在研发伙伴, 为企业推荐可以解决其技术需求的研发伙伴。

关键词: 专利　技术树　TRIZ　研发伙伴
分类号: G306.0

1 引 言

在竞争激烈的市场中, 企业只凭借自身的力量为客户提供产品与服务的方式变得越来越不可能。因此, 在这种情形下, 合作研发对于企业来说越来越重要。而研发伙伴的选择影响着共同研发的成败, 成为了企业重要的战略议题之一。另一方面, 专利文献作为人类最主要的技术载体, 是高度创新的产物, 凝结着整个人类的智慧, 蕴藏着技术前沿的信息。同时, 专利本身也代表着权利, 一个公司专利拥有情况能够充分反映该公司的技术水平, 专利已经成为衡量一个公司技术创新能力与研发能力的重要标准[1]。而且专利对技术内容的描述非常详实, 从中可以掌握许多技术的发明本质, 具有非常高的操作性[2]。

基于以上研究背景, 本文提出一种以专利数据为技术信息载体, 为企业提供潜在研发伙伴的推荐方法。包括以领域专利为数据源, 抽取专利中相关语义特征, 构建领域专利技术树; 以企业具体技术需求文档为依据, 对需求文档中技术领域、技术问题、技术手段与技术效果等关键信息进行抽取, 利用关键信息在领域技术树中的匹配完成相关专利的精准定位; 构建适用于评估基于专利的企业潜在研发伙伴的指标体系, 对相关专利的专利持有人进行评估, 并最终确定针对企业具体技术需求的潜在研发伙伴。

2 国内外研究现状

2.1 技术树构建方法研究

技术树是用来表示在特定技术领域中的产品组件、技术以及技术功能之间关系的树形结构图。目前, 共存在两种构建领域技术树的方法。

(1) 技术树构建方法是围绕着 TRIZ(Theory of Inventive Problem Solving)理论中的技术进化树理论展开。首先分析领域技术发展的现状, 选择进化路线中与本领域最为匹配的技术路线作为模板。然后根据系统优化的要求将现有技术的核心技术特征添加到树形结构中, 作为树形结构的分支; 在主要进化路线的模板上, 匹配本领域技术发展的特征, 以技术路线发展为方向, 掌握和预测本领域技术的发展方向。

(2) 通过对技术、功能等相关信息的抽取, 以原有信息的逻辑结构为基础搭建技术树整体的结构。Yoon 等[3]提出一种基于定量数据的系统技术路线图方法, 通过文本挖掘抽取产品设计与专利中的核心信息, 将其用于现有产品和技术的识别, 从而生成一整套详细的产品和技术配置来支持创新。Lee 等[4]提出一种基于

通讯作者: 郭程, ORCID: 0000-0003-1205-0948, E-mail: guochengchn@163.com。

文献与定性分析的技术路线图构建方法，以宽带网络技术为研究对象，从技术角度构建了该技术分类的技术树。Cascini 等[5]利用计算机辅助专利组合分析系统，从专利中抽取功能相关的语义信息，通过比较发明功能树计算专利的相似度，并以电路中断设备为例从功能角度构建该技术领域的技术树。Fantoni 等[6]以功能–行为–结构(F-B-S)为基础，采用文本挖掘技术，结合关键字、概念、关系与正则表达式实现语义特征的抽取，从专利中抽取功能、行为、结构，然后构建技术树。Choi 等[7]从技术树的产品、技术与功能展示的角度出发，提出一种以 SAO 抽取技术为基础的构建技术树方法。该方法在 SAO 结构抽取之后，通过 K-means 聚类实现 SAO 结构的聚类，然后通过识别单词的特点将聚类后的单词短语类型分为产品、技术、材料与技术属性，同时将 AO 类型分为包含、影响与属性类型，最后结合技术树的三个维度实现领域技术树的构建。Russo 等[8]以 TRIZ 中功能–行为–科学效应–结构本体为基础，提出一种新的功能搜索方法，其将功能定义为技术系统存在的目的，通过将专利中功能、行为、结构等关键信息进行抽取，构建本地知识库，形成领域技术树。王朝霞等[9]认为专利技术方案(Patent Solution, PS)由 SC(技术方案组件)与 SR(技术方案中技术关系)共同组成，并将技术方案组件又细分为 TS(技术系统)、FL(系统流)和 TA(属性)，技术关系又细分为 RS(组成关系)、RA(属性关系)和 RF(功能关系)，最后通过模式匹配实现信息的抽取，结合层次关系形成各模型中元素之间的连接。

综上所述，技术树的第一种构建方法依赖于构建者的知识背景与经验，不容易展开，存在较大的限制；第二种方法现有的研究主要集中于技术树的产品视角、技术视角与功能视角，未考虑功能效果的视角。因此，本文在第二种方法的基础上，以德温特专利数据库为数据源，从专利标题、摘要中抽取专利中所包含的功能、科学效应与功能效果等相关语义信息，结合句法结构，以专利为索引实现领域技术树的构建。

2.2　研发伙伴选择的研究

企业在选择研发伙伴时需要确定拟采用的准则和指标体系，不同的准则体现企业不同的需求。国外学者已经具有一定的研究影响企业选择研发伙伴因素的基础。其中，Cantner 等[10]回顾了选择研发伙伴影响因素相关的研究，以企业资源观为基础，结合德国专利数据，验证了技术重合度、潜在知识流动、未来合作的价值与以往的合作对于选择研发伙伴的正向影响。Lhuillery 等[11]以法国创新调查(CIS)数据为基础，从企业类型、以往合作的经验、知识溢出和知识产权与企业特征 4 方面探索企业在创新项目中合作研发失败的影响因素。Chun 等[12]应用概率模型进行样本选择，以此为依据探索中小型企业研发合作的影响因素。结果表明，内部知识溢出对于中小企业参与研发合作具有正向显著的影响，若不考虑外部知识对于中小企业的重要性，中小企业在建立外部研发合作连接时可能具有劣势。

国内学者在研发合作影响因素研究上也已经有一定的基础。其中，王进富等[13]以 3C 理论为基础，从能力、兼容性、承诺以及合作意向 4 方面考量影响研发伙伴的因素，并将其进一步细分为 9 个方面、35 个影响因素，根据影响因素设计问卷，进行统计分析验证。纪慧生等[14]认为知识特征是企业选择研发伙伴的主要影响因素，合作企业间知识的不匹配以及知识共享效率低下是企业合作研发失败的重要原因，并采用层次分析法构建选择研发伙伴的知识特征指标体系。袁晓东等[15]以专利为数据，提出应用专利信息分析方法来选择企业研发伙伴。该方法将专利分析流程分为技术与主体两个角度，从专利技术发展趋势、专利技术成熟度与核心专利分析三个方面完成技术分析；同时从研发资格比较、相对专利位分析与相对专利优势比较三个方面完成主体分析。宿慧爽等[16]通过梳理选择研发伙伴的相关文献，认为影响企业选择研发伙伴的重要因素包括组织规模、研发强度、互补性资源、组织对外部环境的开放性、相互信任、相容的文化与合作经验等。

另外，选择研发伙伴的方法已经得到众多学者广泛的研究，其中层次分析法、目标规划技术等都在选择研发伙伴中得到成功运用。Song 等[17]提出的基于专利组合评估潜在研发伙伴的方法，克服了选择潜在研发伙伴时集中于企业个人能力评估的弱点。Lee 等[18]提出一种用于大公司选择小公司作为研发伙伴的方法，该方法第一步考虑技术路线作为选择研发伙伴的必要元素，然后使用语义分析产生一系列合适的潜在中小企业研发伙伴，最后使用专利信息的贝叶斯网络

模型实现大公司的潜在中小企业研发伙伴的确定。

综合以上分析发现，选择研发伙伴的方法中，国外学者已经开始使用专利数据对研发伙伴进行识别与评价，而国内在这方面的研究较少。因此，本文拟采用德温特专利数据为数据源，匹配领域专利技术树和企业需求文档，从而更好地辅助企业选择研发伙伴。

3 基于专利的潜在研发伙伴推荐方法设计

本研究利用 TRIZ 理论知识与文本挖掘方法对相关专利的功能、科学效应、技术属性、功能效果等语义特征进行抽取。其中技术属性的抽取用于辅助抽取科学效应，通过各语义特征在专利中的共现关系构建领域专利技术树；然后对企业具体的需求文档进行技术领域、技术问题、技术效果等因素的抽取。通过以上关键信息在领域技术树中的位置对企业需求的目标专利精准定位，挑选出跟企业需求文档最为相关的专利，形成本地需求专利库。企业需求专利获取方案如图 1 所示。

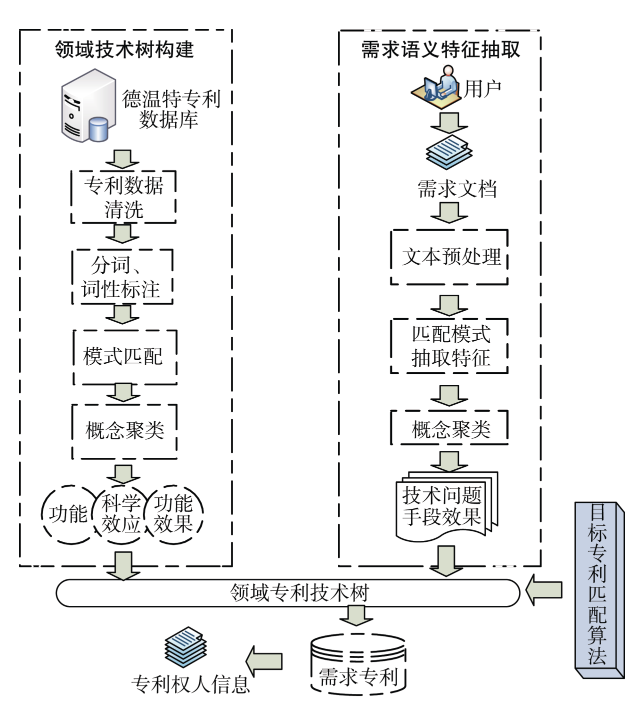


图 1 需求专利获取方案

3.1 专利语义特征抽取方案

结合文献[19]，从专利的文本信息(标题与摘要)中抽取专利的功能、科学效应以及功能效果三部分语义特征。德温特专利数据库中的专利是被专业人员二次改写的，因此其标题与摘要的描述更倾向于标准化。而其摘要也被格式化为几块相对独立的内容，例如 NOV 部分主要描述专利中的核心创新点，包括本专利相较于之前专利最突出的创新之处以及本专利是应用何种技术(科学效应)实现该创新的；而 ADV 部分主要介绍此专利在实现此种技术时所带来的功能效果，如更环保、更安全、更方便等。在 ADV 的描述中有时还会掺杂一些对专利功能的描述，因此此部分内容也应作为专利功能语义特征抽取的源数据；摘要的 USE 部分主要介绍这个专利未来可以应用的技术领域、可被哪些其他技术应用等内容。

因此，根据研究目的，结合以上对德温特专利特点的描述，设计如下的语义特征抽取方案，主要从标题中抽取功能，并从摘要的ADV描述部分辅助完成专利功能的抽取；从摘要的NOV描述部分抽取专利所应用的科学效应；从摘要的ADV描述部分抽取此条专利所达到的功能及其效果。抽取方案如图 2 所示。

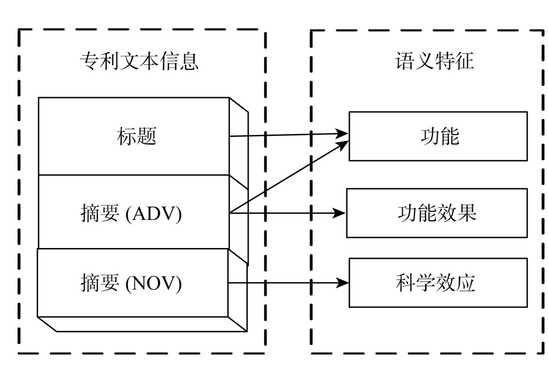


图 2 专利语义特征抽取数据源

3.2 语义特征的标注与抽取

在文本语义特征抽取之前，首先需要对专利的文本特征进行文本预处理。其中包括将文档的文本内容划分为段落，再具体划分到句子，最后对各个句子中的词语进行切分。统计句子中各个词语的词频，结合词语的特征、上下文语境以及文本结构对词语进行词性的标注，同时利用停用词表对无关词进行剔除。文本预处理后，进行语义特征的抽取，分为以下三个部分。

(1) 功能的抽取

功能是对一定环境下用以实现设计意图的输入、输出管理的抽象描述。通俗地说，功能就是物体存在的价值，其所能实现的具体作用。功能表述方法主要

分为三种:

①动词+名词的组合, 如提高温度、减少空间;

②输入、输出变换, 输入、输出可以是能量、物质或信息, 如输入能量为动能, 输出能量为电能;

③行为、状态间的输入、输出变换, 如将液态的水转化为气态的水蒸气。

其中专利中的功能表述最常见的为第一种表述方法。因此, 以专利的标题、摘要中 ADV 部分为专利功能抽取的主要数据源, 以功能的第一种表述为基础抽取规则, 结合功能表述的第二种方法, 通过判断动词之后的名词是否对应为能量、物质或信息的具体表述, 完成专利功能的抽取。

(2) 科学效应的抽取

科学效应是对系统输入、输出间转换过程的描述, 该过程由科学原理和系统属性支配, 并伴有现象产生。本文以常用的 100 条科学效应为基础, 尽可能列举出这些科学效应涉及的名词、动词、副词、形容词, 总体按照机械、热、声、电磁、化学、生物等类别将这些科学效应进行分类, 然后再通过具体科学效应的内容完成小类的识别, 以此为基础, 构建本地基础的科学效应。

以 DII 专利的摘要中 NOV 描述的文本部分为源数据, 识别以上科学效应库中对应的特征词, 识别出专利中应用的科学效应, 为专利进行科学效应的分类。一个专利可能会对应多条科学效应。

(3) 功能效果的抽取

功能效果是对功能的特征描述, 是对功能的一种修饰。专利在通过科学效应实现特定功能的同时必然带来相关功能效果的实现, 如提高安全性、提高可靠性、快速等。功能效果是对实现具体功能的一种额外效果的描述。它在专利的文本表示中有两种表达方式, 第一种是以修饰表示功能动词的副词形式存在的, 如快速杀菌、安全消毒等, 第二种是通过动词+名词的形式来表达功能效果的, 如提高安全性、提高效率等。其中, 专利中最常见的表达方式为第二种。功能效果的第二种表述与功能的第一种表述的唯一不同之处在于功能的动词后接的名词为能量、物质、信息等表示词, 而功能效果动词后接的名词多为像安全性、效率、稳定性等类似的词。因此, 在从摘要 ADV 部分中抽取功能效果时, 通过以上两种模式来识别。

3.3　基于专利语义特征的技术树构建

(1) 对从专利中抽取的语义特征进行特征内部层次结构的确定, 通过特征概念的同义合并, 实现概念内部相关信息的聚类, 形成各个语义特征内部相对独立的基础语义特征知识块;

(2) 以同一条专利中, “专利功能”与“专利科学效应”、“专利科学效应”与“专利功能效果”之间的共现关系为标准分别实现领域技术树中“功能”概念与“科学效应”概念、“科学效应”概念与“功能效果”概念的联接, 构建领域技术树的整体框架;

(3) 将专利数据的著录项目, 如专利号、专利权人、被引数量等特征添加到领域技术树的叶节点中, 实现领域技术树从专利著录项特征到专利语义特征的全方位表示, 从多维度全面展现领域技术中的知识以及知识之间的关联。最终的技术树结构如图 3 所示。

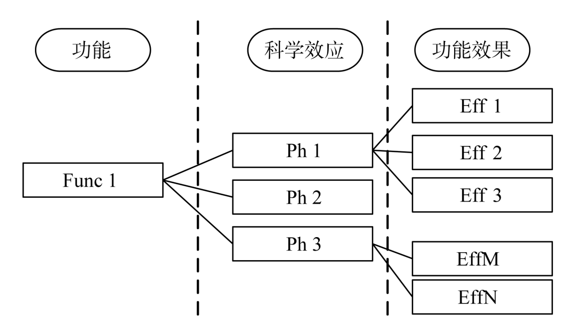


图 3　领域语义特征技术树的结构

3.4　企业具体需求的语义信息抽取

将需求做成一个结构化的表格交由需求企业填写, 包括标题、需求描述、需求背景、可能的技术方向、不考虑的技术方案等条目。本研究主要从这些条目中进行信息的抽取。其中, 标题、需求描述与描述背景主要通过动宾结构的方式匹配技术效果与技术问题, 可能的技术方向和不考虑的技术方案因其本身字段较简练, 因此可以直接作为技术手段字段进行抽取。

技术问题是需求方迫切需解决的问题, 即未来的技术方案能够实现的功能, 因此对技术问题的抽取就等同于对需求描述中功能特征的抽取。汉语中对功能的描述也多为动词与名词的组合, 而且动词前可能出现副词, 名词前可能出现形容词, 如: (副词)+动词+(形容词)+名词。其中, 括号中的副词与形容词可以省略。如快速地消除水垢、安全地提供氧气。本文总结需求方所提供的文本信息中涉及功能的动词, 如表 1 所示。

表 1　需求文档中功能特征词

编号	动词	同义词	后续可能名词
1	吸收	吸附	能量、物质
2	积累	积攒、累积、积聚	能量、物质
3	集中	浓缩	能量、物质
4	检测	检查、测量	能量、参数、物质
5	避免	防止	能量、物质
6	产生	生产	能量、物质

技术效果是实现特定功能时必然会带来的效果，功能的表述模式中，副词更多的是表示技术的效果，如上文提到的：安全地、快速地。因此需求文档中的技术效果可以通过识别修饰功能动词的副词来完成。

3.5　领域专利技术树中需求专利的匹配算法

领域专利技术树已将本领域中所有科学效应实现的功能以及达到的功能效果进行全面、直观地展示，并且通过索引专利号已将哪些专利应用哪些科学效应、实现哪些功能、达到哪些功能效果进行了标注。因此，本阶段最直接的匹配方法就是首先通过企业具体需求中抽取的技术领域与技术问题对应技术树中的技术领域与功能字段，满足需求方对功能的基本需求；其次将需求文档中抽取的技术手段与技术效果对应于领域专利技术树中的科学效应与功能效果。匹配方法如图 4 所示。

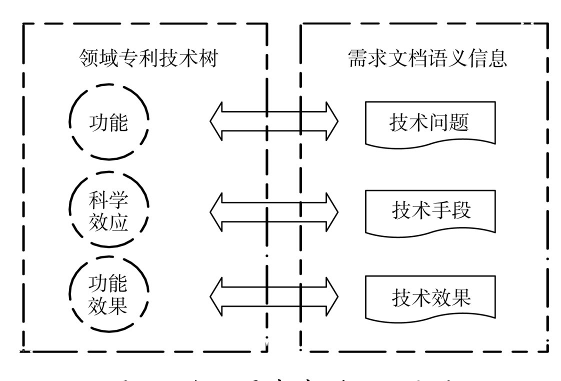


图 4　企业需求专利匹配方法

4　基于专利的企业潜在研发伙伴的评估

企业在寻找研发伙伴的过程中需要考虑的信息的主要包括：潜在研发伙伴的技术特征，如技术实力、研发实力；潜在研发伙伴的开放特征，如研发开放性、与需求企业合作程度等；潜在研发伙伴的合作研发效果特征，包括历史研发合作研发效果的评价。

评价有效合作研发的标准，首先通过专利分析对潜在研发伙伴进行评估。本文所提出的基于专利的企业潜在研发伙伴分析，主要是由三大类组成：技术实力、研发开放性以及合作研发效果，如图 5 所示。

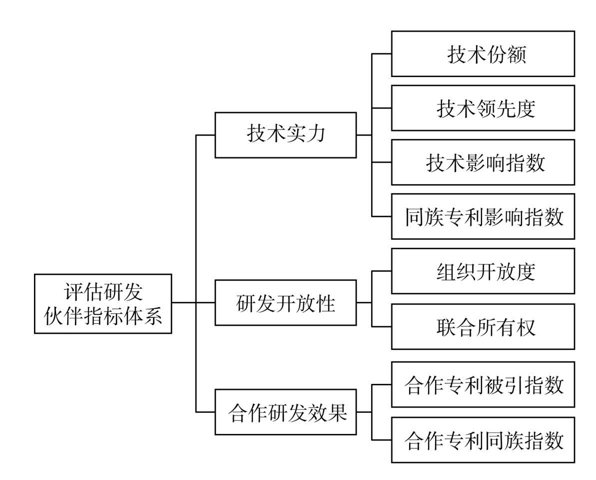


图 5　评估潜在研发伙伴的指标体系

技术实力和研发开放性着眼于研发伙伴的特点。技术实力代表潜在研发伙伴在合作领域的技术能力，包括技术、操作的知识和经验。潜在研发伙伴的技术实力根据 4 个子标准进行评估：技术份额、技术领先度、技术影响指数和同族专利影响指数。研发开放性通过组织开放度和联合所有权两项子指标测量得到。预期合作研发效果，包括合作专利被引指数和合作专利同族指数两个子标准。各子指标的计算方法如表 2 所示。

表 2　识别潜在研发伙伴子指标计算方法一览表

指标	计算方法
技术份额	潜在研发伙伴在某一技术领域的专利申请量/该技术领域所有企业专利申请量
技术领先度	潜在研发伙伴某一技术领域专利平均被引数/该领域所有专利平均被引数
技术影响指数	某一技术领域专利被引次数前10%的专利中，潜在研发伙伴(机构、企业)专利所占的比例
同族专利影响指数	潜在研发伙伴在某一技术领域内平均专利同族数/该技术领域内平均专利同族数
组织开放度	某一技术领域潜在研发伙伴具有共享所有权专利数/该技术领域内潜在研发伙伴的专利总数
联合所有权	某一技术领域需求企业与潜在研发伙伴共享所有权的专利数/该技术领域内需求企业所有具有共享所有权性质的专利数
合作专利被引指数	某一技术领域潜在研发伙伴共享所有权专利的平均被引数/该技术领域内潜在研发伙伴所有专利的平均被引数
合作专利同族指数	某一技术领域潜在研发伙伴共享所有权专利的平均同族数/该技术领域内潜在研发伙伴所有专利的平均同族数

5　实验及结果分析

5.1　数据收集

本文以热水器防水垢技术领域作为研究对象，该技术领域主要涉及热水器、水垢、防水垢、除水垢等关键词，根据关键词确定该技术领域的检索式为 TS=(((scale* and (prevention* or prevent* or inhibit* or control* or avoid* or removal* or reduc*)) or antiscal* or scaleinhibition or anti-scal*) and water* and heater*)，检索年限为 1963 年 1 月 1 日到 2015 年 12 月 31 日，在德温特数据库中共检索到 1 909 条专利。

5.2　专利语义特征抽取

(1) 从专利中标注和抽取专利的功能、专利所应用的科学效应以及专利达到的功能效果这三种语义特征。其中专利的功能以标题与摘要的 ADV 部分为数据源进行抽取，科学效应以专利摘要的 NOV 部分为数据源，功能效果以摘要的 ADV 部分为数据源。专利语义特征抽取结果如表 3 所示。

表 3　专利语义特征抽取举例

专利号	专利功能	专利科学效应	专利功能效果
JP2015021723-A	without forming water scales; heat the water	electromagnetic induction	achieving energy-saving effect; safe manner
CN105402894-A	risk of electric shock is avoided; formation of scale is avoided	electro-thermal conversion	electro-thermal conversion efficiency; extending the use of water heaters life
CN105241095-A	eliminates scale deposit	spring and needle action (vibration)	Ensures safety, Saves energy
WO2016006225-A1	anti-scale	scale inhibitor	extended life, easy, economical
KR2015057208-A	generation of the scale is suppressed	magnetic force	security of the water purifier is improved
JP2014238200-A	preventing the adhesion of scale	control temperature	energy saving is possible performing highly efficient heat exchange maintenance operation is performed easily

(2) 专利中抽取的功能多为防止/抑制水垢生成，少数专利能够实现清除水垢的功能。抽取的专利所应用的科学效应经过合并共分为 11 种，分别为磁场(Magnetic field)、超声波(Ultrasound)、电场(Electric field)、振动(Vibration)、气体/蒸汽(Gas/steam)、相变(Phase change)、循环流动(Circulation flow)、温度控制(Control temperature)、压强控制(Control pressure)、功能性陶瓷(Functional ceramics)与阻垢剂(Scale inhibitor)。其中属于化学科学效应的只有阻垢剂(Scale inhibitor)，而其他均属于物理科学效应。

5.3　领域专利技术树构建

以专利为索引，根据各项语义特征在专利中的共现关系以及上下文语境关系来完成各项语义特征之间的连接关系。汇总热水器防水垢领域专利抽取的语义特征，对其语义特征进行同义合并，实现该技术领域专利技术树的构建。其中，该领域技术树的部分内容，如图 6 所示。

由图 6 可知，热水器防水垢技术领域功能可以分为抑制水垢与除水垢两种。其中振动科学效应既可以应用于抑制水垢也可以应用于除水垢。而抑制水垢还可以通过磁、陶瓷、控制温度、阻垢剂(化学)等方法实现，可以达到增加安全性、快速、方便、高性能等功能效果。

由于专利的语义特征本身来源于专利的文本信息，而每个专利都具有其自身的索引——专利号，并且专利除了语义信息，还带有一些著录项目，如专利的所有者(专利权人)、专利申请/公开的时间以及专利在哪些国家/地区具有专利权等相关信息。因此在构建领域技术树后，还需将专利相关的著录项目添加到技术树中，形成完整的专利维度，全方位刻画每一条专利。添加专利著录项目后技术树的部分如图 7 所示。

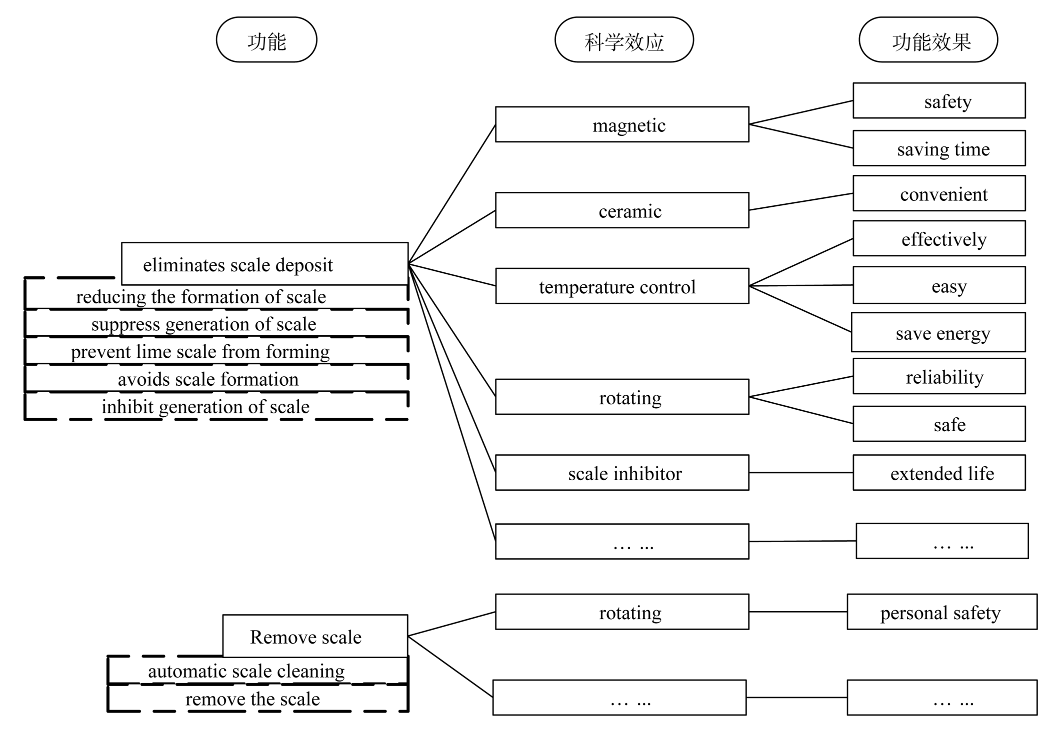


图 6　热水器防水垢领域技术树举例

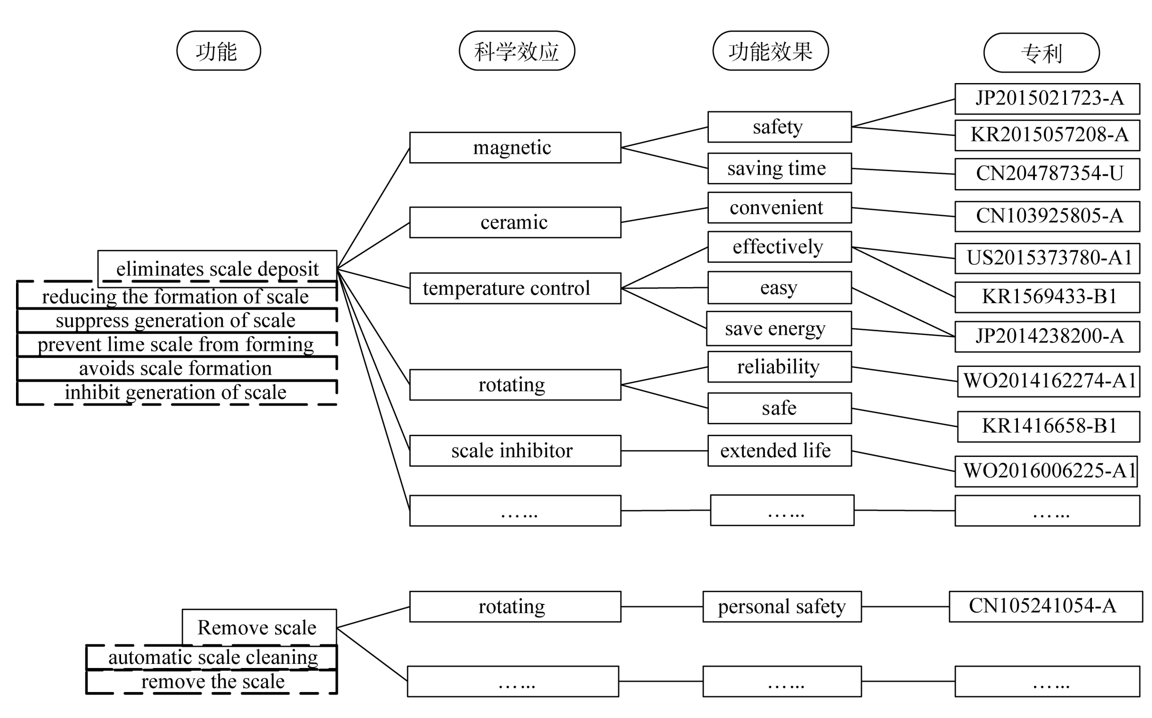


图 7　热水器防水垢领域技术树添加专利号举例

图 7 可以全方位的展示专利的语义信息和专利的时间、地点等信息，如专利号为 JP2015021723-A 的专利是应用磁科学效应实现抑制水垢生成的功能，同时达到了提高安全性的效果，又可以展示出其自身属于在日本申请的专利，公开年份为 2015 年。因此，添加专利著录项目的技术树展现的信息更加全面。

5.4　需求关键信息的抽取

本研究以 A 公司作为该实证研究的需求企业。该公司对于热水器防水垢技术的部分需求描述如表 4 所示。

表 4　热水器防水垢技术需求文档部分内容

需求项目名称	需求项目具体内容
需求标题	寻求电热水器不结垢或结垢后可快速清洗的解决方案
需求描述	1. 储水式电热水器使用一段时间后，内部结水垢、沉积一些污垢；长时间不用，担心内胆里的水滋生细菌，洗澡时不放心。 2. 内胆里面脏了以后，用户不知道如何清洗，不会/不能自行清洗。 水垢成因：……
需求背景	寻找如下两种解决方案： 1. 能够减轻热水器内胆结垢的方案； 2. 结垢后能够方便地进行水垢清理的技术方案。
可能的技术方向	暂无技术领域限制
领域	个人健康，净化
标签	电热水器、除垢、防垢、清垢

以表 4 中企业具体需求中相关项的具体内容为企业需求的源数据，对该数据进行相关语义信息的抽取。其中，技术领域直接从企业需求文档的“领域”条目结合标签完成该需求文档技术领域的定位。因此，本企业需求文档的技术领域为热水器、个人健康、净化。技术问题通过识别需求文档中动宾结构，挑选出其中能够代表功能需求的有意义的动宾结构，如结水垢、滋生细菌、自行处理等。技术手段以识别名词为主，结合技术手段语料库识别需求文档对于技术手段的限定，如超声波、磁方法等。技术效果通过识别修饰动词结构的副词或识别表示效果的动宾结构来实现，如方便水垢清理、快速清洗等。综上，企业需求文档的语义特征汇总如表 5 所示。

表 5　企业需求文档中的语义信息抽取结果

需求语义特征	需求项目语义特征内容
技术领域	热水器、个人健康、净化
技术问题	减轻热水器内胆结垢、沉积污垢、水垢不能自行清理、滋生细菌
技术手段	超声波、磁方法
技术效果	方便，性能，快速，成本

5.5　潜在研发合作伙伴的识别及评估

通过匹配算法，结合本实证中构建的领域专利技术树与企业需求文档中抽取的语义特征，实现热水器防水垢技术领域需求专利的匹配。

首先通过企业需求文档中抽取的技术问题与领域专利技术树中功能项目的匹配实现专利集合的精炼，将需求文档中描述的减轻结垢、清理水垢、滋生细菌等技术问题与专利功能进行匹配，共筛选出 1 219 条相关专利能够解决相应的技术问题。然后再通过企业需求文档中抽取的技术效果的语义特征方便、性能、快速与成本等与领域专利技术树中功能效果项目的匹配对相关专利集合再进行精炼。其中，在能解决以上技术问题的基础上，又能达到相应技术效果的专利共有 932 条。对这 932 条相关专利所应用的科学效应进行分析，结果仍包含所有抽取的 11 种科学效应。

分析 TRIZ 中提出的八大进化法则，结合热水器防水垢技术领域抽取的科学效应，选取合适的进化法则作为筛选先进技术专利的基础。本实例选取第七条进化法则即向微观级与场的应用进化法则。根据该进化法则的要求，结合各科学效应在热水器防水垢技术领域应用专利出现的时间分布，本文选择了磁场(Magnetic field)、超声波(Ultrasound)、电场(Electric field)、温度控制(Control temperature)等相关科学效应作为先进技术，将应用这些科学效应的专利视为先进专利，该专利集共包含 612 条相关专利，以此专利集合的专利权人为潜在研发伙伴。

根据企业潜在研发伙伴的评估指标体系，分别对主要专利权人进行各个指标的打分，得出各个专利权人的总得分，松下公司的各指标得分如表 6 所示。

根据对潜在研发伙伴评价指标体系中各指标值的计算，再对专利数排名前 5 名的专利权人(潜在研发伙伴)的指标值进行计算，并对结果进行归一化，结果如表 7 所示。

将各潜在研发伙伴指标值归一化的结果求和，计算出各潜在研发伙伴的综合评价值，对其潜在研发伙伴的综合评价值进行排序，得出潜在研发伙伴的排序，如表 8 所示。

表 6　松下公司各评估指标得分

编号	指标名称	计算要素 1	计算要素 2	计算要素 3	计算要素 4	结果
1	技术份额	总专利数 612	松下专利数 19			0.03
2	技术领先	总被引数 2 987	平均被引数 4.88	松下总被引数 103	松下平均被引数 5.42	1.11
3	技术影响指数	前 10%专利数 61	松下占专利数 9			0.14
4	同族专利指数	总同族数 1 774	平均同族数 2.89	松下总同族数 60	松下平均同族数 3.15	1.04
5	机构开放性	松下共享所有权专利数 10	松下专利总数 19			0.52
6	联合所有权	共享专利数 0	总共享专利数 9			0
7	合作专利被引指数	共享专利被引总数 61	共享专利平均被引数 6.1	松下平均被引数 5.42		1.13
8	合作专利同族指数	共享专利同族总数 59	共享专利平均同族数 5.9	松下平均同族数 3.15		1.87

表 7　潜在研发伙伴指标值归一化结果

指标名	大金工业	松下	三菱电机	夏普	飞利浦
技术份额	1	0.9	0.8	0.4	0.2
技术领先度	0.19	0.475	0.584	1	0
技术影响指数	0.25	0.5	1	0.5	0
同族专利影响指数	0.783	0.569	0.738	1	0.719
组织开放性	0.789	0.292	1	0.684	0
联合所有权	0	0	0	0	0
合作专利被引指数	1	0.787	0.322	0.189	0
合作专利同族指数	0.752	0.135	0.519	1	0

表 8　潜在研发伙伴综合评价值排序结果

公司名	综合评价值
三菱电机	4.963
夏普	4.773
大金工业	4.764
松下	3.658
飞利浦	0.919

6　结　语

综上所述，本文对特定技术领域相关专利的功能、科学效应与功能效果等语义特征进行抽取，构建领域专利技术树；针对企业具体需求，对需求中技术领域、技术问题、技术手段与技术效果等关键信息进行抽取；通过将科学效应匹配到TRIZ进化法则中，定位先进的科学效应，结合需求中描述的技术问题与技术效果等其他信息对先进技术的相关专利进行精准定位，找出潜在的合作伙伴；并且从潜在研发伙伴的技术实力、研发开放度以及合作研发效果三个方面入手，设计评价指标，构建基于专利的企业潜在研发伙伴的评估指标体系，对研发伙伴进行评估。

同时，本文还存在一些需要改进的方面。在语义特征抽取的过程中，因为中文语言存在较大的灵活性、表述形式多样，因此抽取模式的准确率还有待进一步提高。未来研究可以在对大量中文文档深入分析的基础上，定义更为准确的模式，提高关键信息抽取的准确性。

参考文献：

[1] 包昌火，谢新洲. 竞争情报与企业竞争力[M]. 北京：华夏出版社，2001：137-145. (Bao Changhuo, Xie Xinzhou. Competitive Intelligence and Enterprise Competitive Power [M]. Beijing: Huaxia Publishing House, 2001: 137-145.)

[2] 谢炜. 中国专利产出研究[D]. 成都：电子科技大学，2005. (Xie Wei. Research on China's Patent Output[D]. Cheng Du: University of Electronic Science and Technology of China, 2005.)

[3] Yoon B, Phaal R, Probert D. Morphology Analysis for Technology Roadmapping: Application of Text Mining [J]. R&d Management, 2008, 38(1): 51-68.

[4] Lee S, Kang S, Park E, et al. Applying Technology Roadmaps in Project Selection and Planning[J]. International Journal of Quality & Reliability Management, 2008, 25(1): 39-51.

[5] Cascini G, Zini M. Measuring Patent Similarity by Comparing Inventions Functional Trees [A]//Computer-aided Innovation (CAI)[M]. New York: Springer US, 2008: 31-42.

[6] Fantoni G, Apreda R, Dell'Orletta F, et al. Automatic Extraction of Function–Behaviour–State Information from Patents [J]. Advanced Engineering Informatics, 2013, 27(3): 317-334.

[7] Choi S, Park H, Kang D, et al. An SAO-based Text Mining

Approach to Building a Technology Tree for Technology Planning[J]. Expert Systems with Applications An International Journal, 2012, 39(13): 11443-11455.

[8] Russo D, Montecchi T, Liu Y. Functional-based Search for Patent Technology Transfer[C]//Proceedings of the 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Chicago: American Society of Mechanical Engineers, 2012: 529-539.

[9] 王朝霞, 邱清盈, 冯培恩, 等. 机械产品专利技术方案信息抽取方法[J]. 机械工程学报, 2009, 45(10): 198-206. (Wang Zhaoxia, Qiu Qingying, Feng Peien, et al. Information Extraction Method of Technical Solution from Mechanical Product Patent[J]. Journal of Mechanical Engineering, 2009, 45(10): 198-206.)

[10] Cantner U, Meder A. Technological Proximity and the Choice of Cooperation Partner [J]. Journal of Economic Interaction and Coordination, 2007, 2(1): 45-65.

[11] Lhuillery S, Pfister E. R&D Cooperation and Failures in Innovation Projects: Empirical Evidence from French CIS Data[J]. Research Policy, 2009, 38(1): 45-57.

[12] Chun H, Mun S B. Determinants of R&D Cooperation in Small and Medium-sized Enterprises [J]. Small Business Economics, 2012, 39(2): 419-436.

[13] 王进富, 魏珍, 刘江南, 等. 以企业为主体的产学研战略联盟研发伙伴选择影响因素研究——基于 3C 理论视角[J]. 预测, 2013, 32(4): 74-80. (Wang Jinfu, Wei Zhen, Liu Jiangnan, et al. Research on Influencing Factors of the IUR Strategic Alliance's R&D Partner Selection—Based on the Perspective of 3C Theory [J]. Forecasting. 2014, 32(4): 74-80.)

[14] 纪慧生, 王红卫, 陆强. 基于知识特征的企业研发伙伴选择[J]. 沈阳工业大学学报: 社会科学版, 2011, 4(2): 137-140. (Ji Huisheng, Wang Hongwei, Lu Qiang. R&D partner Selection of Enterprise Based on Knowledge Characteristics [J]. Journal of Shenyang University of Technology: Social Science Edition, 2011, 4(2): 137-140.)

[15] 袁晓东, 陈静. 专利信息分析在技术创新合作伙伴选择中的应用[J]. 情报杂志, 2011, 30(8): 22-27. (Yuan Xiaodong, Chen Jing. The Application of Patent Information Analysis Method in the Choice of Cooperative Technological Innovation Partners [J]. Journal of Intelligence, 2011, 30(8): 22-27.)

[16] 宿慧爽, 兰衍霏, 衣兰文. 企业研发合作伙伴选择研究综述: 基于影响因素的视角[J]. 现代管理科学, 2013 (6): 48-50. (Su Huishuang, Lan Yanfei, Yi Lanwen. Review on Selecting Enterprises R&D Partners: Based on the Perspective of Influence Factors [J]. Modern Management Science, 2013 (6): 48-50.)

[17] Song B, Seol H, Park Y. A Patent Portfolio-based Approach for Assessing Potential R&D Partners: An Application of the Shapley Value [J]. Technological Forecasting and Social Change, 2016, 103: 156-165.

[18] Lee K, Yoon B. A Method for Partner Selection in R&D Collaboration Between Large Companies and SMEs Using Patent Information[C]//Proceedings of the 2013 Technology Management in the IT-Driven Services (PICMET). 2013: 1886-1891.

[19] 翟东升, 夏军, 张杰, 等. 基于专利特征抽取的技术树构建方法研究[J]. 情报学报, 2015, 34(7): 717-724. (Zhai Dongsheng, Xia Jun, Zhang Jie, et al. Research on Construction of Technology Tree Based on Patent Feature Extraction [J]. Journal of the China Society for Scientific and Technical Information, 2015, 34(7): 717-724.)

作者贡献声明:

翟东升: 提出研究选题及思路, 论文最终版本修订;
翟东升, 郭程, 张杰: 设计研究方案;
郭程, 夏军: 处理数据, 实现推荐方法, 完成实验, 论文起草。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据见期刊网络版 http://www.infotech.ac.cn。

[1] 翟东升, 郭程. Patent1909.xlsx.热水器防水垢技术专利数据.

收稿日期: 2016-08-29
收修改稿日期: 2016-10-27

Recommending Potential R&D Partners Based on Patents

Zhai Dongsheng Guo Cheng Zhang Jie Xia Jun
(School of Economics and Management, Beijing University of Technology, Beijing 100024, China)

Abstract: [**Objective**] This study presents a recommendation method based on patents to accurately identify potential R&D partners. [**Methods**] First, we extracted the functions, scientific impacts and functional effects of the related patents based on the TRIZ theory. Second, we constructed a patent technology tree, which was mapped with key information from the enterprise needs. Finally, we identified and evaluated the potential R&D partners in accordance with the patentee. [**Results**] We successfully assessed the retrieved R&D partners with the proposed method based on water heater related patents. [**Limitations**] The accuracy of semantic feature extraction needs to be improved. [**Conclusions**] The proposed method could find and evaluate the potential R&D partners for enterprises effectively.

Keywords: Patent Technology Tree TRIZ R&D Partners

HighWire 与 Hypothesis 合作整合出版物的注释

近日, HighWire Press 和 Hypothesis 宣布双方达成合作伙伴关系, 计划为 HighWire JCore 平台上发布的 3 000 多种期刊、图书、参考读物和会议论文集增加开放注释功能。

HighWire JCore 平台上的出版商可以在其出版物中实施和控制自己的注释的层级和可见度。注释是研究人员和学者的基础性活动。

到目前为止, 期刊的注释解决方案受到了很多限制, 很大程度上限制了其实用性。随着基于标准的、开源的、可互操作的注释范式的出现, 这一情况开始改变。

"HighWire 与合作伙伴共同创建了一个创新的、开放的平台解决方案, 能够改善科研情况, 为读者带来更多价值。自 2015 年以来, 我们一直在支持'知识注解联盟'。" HighWire Press 首席执行官 Dan Filby 表示, "与 JCore 集成的开放注释技术支持更快的研究工作流程和更高效的沟通。我们期待与希望使用 Hypothesis 来改善其内容和工作流程的交互性的出版商合作。"

Hypothesis 是一个非营利性的注释技术组织, 成立于 2011 年, 与出版商、教育工作者、研究人员和记者合作, 旨在实现互联网注释。在科研环境中的使用案例包括: 出版前同行评审、出版后注释和社区审查; 作者对自己作品进行的说明, 包括对以前文章的更新、邀请的讨论、增强的脚注、更正和勘误等。70 多家大型出版商、平台和技术组织齐聚一堂, 支持"知识注解联盟"的互操作愿景。

(编译自: http://home.highwire.org/sites/default/files/HighWire%20Hypothesis%20PR%20AM%20Comments%20%281%29.pdf)

(本刊讯)